Prediction of habitat suitability for Wild Goat, *Capra aegagrus* (Erxleben, 1777) in Kavir National Park, Semnan Province, Iran

Shadloo, Sh. and Naderi, S[*].

Department of Environment, Faculty of Natural Resources, University of Guilan, Guilan, Iran

Habitat with its components is the most important need of wildlife species. So that, habitat management is a critical procedure in wildlife conservation. In the present study, desirable habitats of Kavir National Park for Wild goat are determined in two distinguished warm and cold seasons, applying Random Forests, Maximum Entropy, and Classification Tree algorithmsa and for this matter "Salford Predictive Modeler" and MaxEnt software were used. Consequently, an ensemble habitat suitability map was provided based on weight-averaging method. The results illustrated that elevation, water resources, vegetation type, and slope are the most important factors in predicting Wild goat's suitable habitats with a slightly different impact on each season. In general, among the existing plant species, Wild goat prefers *Artemisia sieberi*, *Stipa arabica*, *Zyghophylom eurpterum*, and *Sieditzia rozmarinus*. Also, the altitudes higher than 1050 meters above sea level, slopes over 20 to 35 percent, and distances less than 5,000 meters to water resources are desirable. Meanwhile, several suitable patches for Wild goats were found which are greatly in need of conservation. Considering the importance of water resources in this region, in order to improve the quality of the habitat, further study on water resources status is suggested.

Key words: *Wild goat, Kavir National Park, habitat suitability, MaxEnt, random forest, classification and regression tree.*

INTRODUCTION

Wild goat (*Capra aegagrus*), taxonomically is included in Order Artiodactyla and Family Bovidae. A study by Naderi *et al.* (2008) demonstrated that *Capra aegagrus* is one of the first artiodactyl species being domesticated in more than 10,000 years ago in the Fertile Crescent and is the ancestor of domestic goat. This species is of high economic and ecological importance, meanwhile, is listed as vulnerable in the IUCN Red List (Weinberg, 2008). This species is considered an important prey for leopard and due to the status of leopard, being endangered, assessing the Wild goat's habitat is crucial for management and conservation processes in different views.

Today researchers use different species distribution models to conclude the presence and location of different species. The models can be the basis of species conservation and decision-making for regional development (Fook *et al.*, 2009).

One of the main reasons of wildlife being on the wane, is habitat destruction, in a way that many species are forced to live in isolated and semi-isolated populations in fragmented patches

(Hirzel & Guisan, 2002). Hence, it is advisable to determine suitable habitats and to act accordingly. Nonetheless, in some cases, especially when encountering large-scale regions, there are some difficulties such as limited time and narrow budgets that make such studies difficult, if not impossible, to take place. Therefore, scientists developed some methods to facilitate such studies, and currently, these methods are widely being used in wildlife management (Peterson & Vargas, 1993). These types of models examine the quantitative and qualitative relationship between the presence of species and environmental variables because the quality of a habitat results from interactions between biological and non-biological variables (Kocev *et al.*, 2010).

In various regions, the crucial factors do not affect habitat suitability for Wild goat, similarly. Several studies on Wild goat's suitable habitats in Iran were carried out using logistic regression, Ecological Niche Factor Analysis (ENFA), and Maximum Entropy (MaxEnt) and showed that topographic variables, water, and human resources, and vegetation cover were the most important factors for habitat selection by a Wild goat (Hosseini *et al.*, 2016; Naderi *et al.*, 2014; Sarhangzadeh *et al.*, 2013; Shams Esfandabad *et al.*, 2010). In a study in alpine habitats, Colorado, habitat models of mountain goat (*Oreamnos americanus*) were developed applying the logistic regression method. The results demonstrated that mountain goats prefer areas on moderate slopes, near escape terrain, at middle-elevations, and on southerly exposures (Gross *et al.*, 2002). In another study, in southwestern coastal British Columbia, the results indicated that both sexes, especially males, mostly prefer mature and old forests (Taylor *et al.*, 2006).

The purpose of this research is to evaluate the habitat suitability of *Capra aegagrus* in Kavir National Park, based on the presence data of this species. We categorized a year into two warm (the first six months of the year) and cold (the second months of the year) seasons. We attempted to ascertain the most essential parameters and to highlight the suitable areas for the species in Kavir National Park for the first time, by using the ensemble modeling technique. It is noteworthy that neither "Salford Predictive Modeler" (SPM) software nor an ensemble technique has ever been used to determine this species habitat in Iran. In the end, some suggestions are presented for conservation purposes.

MATERIAL AND METHODS
Study area
This study has been done in Kavir National Park, located in central plateau of Iran (Fig. 1). Kavir, with an area of about 4200 square kilometers, is the second largest National Park in Iran, after Urmia Lake (Atarodi, 2010). This area is located in latitude 51 degrees 25 minutes to 53 degrees and 5 minutes E, longitude 34 degrees and 20 minutes to 35 degrees and 10 minutes N.

Habitat suitability modelling
In this study, three methods based on machine learning have been used. Machine learning, especially predictive models, is increasingly being used in the structure of ecological models (Džeroski, 2001). The three methods used in this project are Classification Tree Algorithms (CTA) (Breiman *et al.*, 1984), Random Forest (RF) (Breiman, 2001), and MaxEnt (Phillips *et al.* 2006; Elith *et al.* 2010).

Sampling and field studies
Collecting data started from August 2017 and ended in July 2018. Indices such as horn, fur, teeth, feces, carcasses, etc. as well as direct observations of the species were recorded by Global Positioning System (GPS), as points of presence. Sampling was carried out in different environmental strata of the region (Fig. 2).
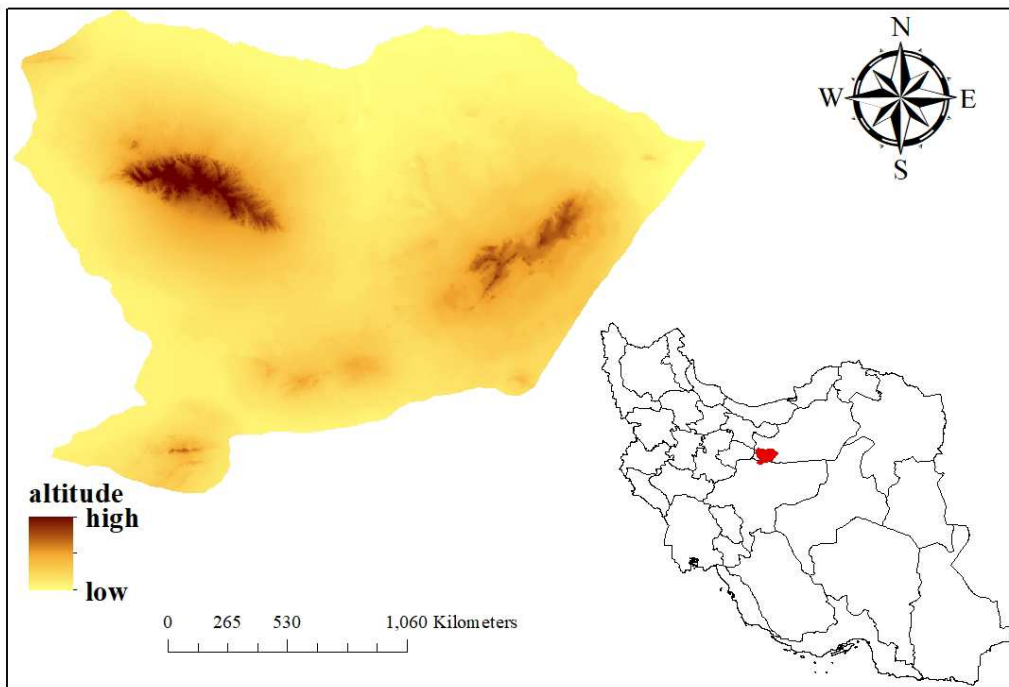
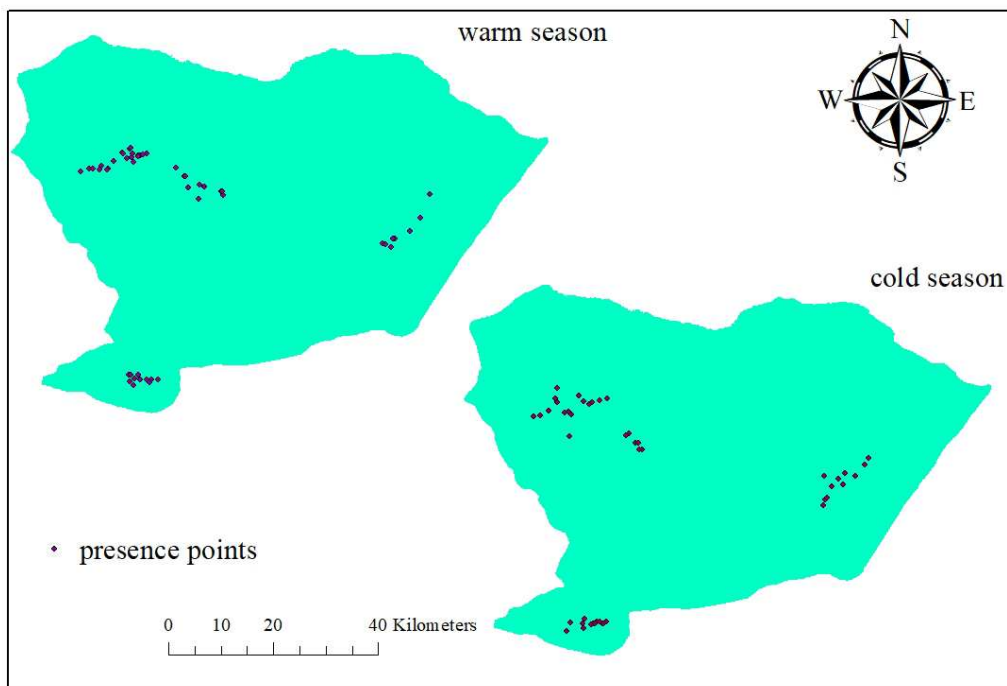**FIGURE 1.** The studied area in Kavir National Park, Semnan Province.

FIGURE 2. Presence records of Wild Goat in Kavir National Park.

Bias in datasets can have a significant influence on results. If modeling is performed with bias, it will be better to be careful when interpreting the results (Leitão *et al.*, 2011). There are two common ways to reduce the impact of bias: 1) refining some of the points of presence and 2) defining the

background points (pseudo-absence points). The results indicated that the former is more appropriate if the number of dataset is big enough to remove some data (Kramer-Schadt, 2013). Hence, in present study, the presence points were refined through a spatial filtration by keeping one occurrence point within a radius of 1 kilometer. This process was done via "Spatially Rarify Occurrence Data" command in SDMtoolbox, ArcMap 10.5 software (Brown, 2014).

Sample size

In a study, it was shown that in machine learning models, with only 10 sample points, the accuracy of the model, was at 90% of its highest accuracy, and with 50 samples it reaches near to the highest accuracy (Stockwell & Peterson, 2002). Figure 2 illustrates the refined presence points in both warm and cold seasons. The numbers of presence records in the region are 52 in the warm season (the six warm months) and 46 in the cold season (the six cold months).

Habitat variables

According to behavioral and ecological characteristics of Capra aegagrus, previous literature and the characteristics of the study area, the potential variables in affecting Wild goats distribution were chosen as follows: topographical variables including altitude, slope, and aspect, vegetation cover variables (density and type), distance to water resources, anthropological variables such as distance to roads and DoE guard stations, and climate variables containing monthly minimum and maximum average temperature, each for cold and warm seasons respectively, and monthly average rainfall.

For this purpose, some data were collated from Kavir National Park Bureau and Meteorological Organization. There are no meteorological and weather stations in Kavir National Park, so data from six stations outside the region were interpolated (Childs, 2004) to obtain continuous values of climatic variables inside the park. Interpolation was carried out using the 'IDW' method in ArcMap version 10.5 and climatic layers were produced. According to Khosravi *et al.* (2014), the optimum interpolation methods for obtaining average temperature and precipitation are Co-kriging and IDW, respectively. Though, when the number of observation points (meteorological stations) is less than 30, is highly recommended to use the IDW method (Sluiter, 2009; Perry & Hollis, 2005).

To build the vegetation density map, two images of Landsat 8 for the months of May and January, due to the lack of clouds in these months' images and respectively their lowest and highest density among other months, were downloaded from www.earthexplorer.usgs.gov, with a cell size of 30 × 30 meters. Then, using ArcMap 10.5, this image turned into NDVI Index, to represent the vegetation density map of warm and cold seasons. The rest of the layers were produced in the same cell size as 30 × 30.

To determine the correlation between environmental variables, Principal Component Analysis (PCA) in ArcMap 10.5 was used. Consequently, among the variables with a high correlation (above 0.75), only one variable was selected to participate in modeling. In this process, qualitative and quantitative variables were investigated for both datasets related to warm and cold seasons.

Modelling

We prepared data using R 3.5.0 to run CTA (Salford, 2013) and RF (Salford, 2016), with SPM software. 'Cross-validation' method was chosen to evaluate both MaxEnt and CART models and also, 'out of the bag', was selected to calculate the AUC value of the models. Then SPM software outputs were turned into lattice files (Mi *et al.*, 2017), with a neighboring distance of 30 meters (the cell size of the layers); and the files were scored and turned into raster layers.

'Permutation' method was defined to calculate variables importance in RF. According to Breiman and Cutler (2007) permutation is a standard approach to compute variable importance. RF models can be biased in such a way that categorical variables with a large number of levels are more likely to be chosen as a splitter in trees. Permutation importance method is helpful to overcome this issue and to decide the significance of variables (Altmann *et al.*, 2010).

Better well-established models can be concluded if ensemble forecasts are produced (Araújo & New, 2007). The suitability habitat map of Wild goat in Kavir National Park was provided in an ensemble form, based on averaging the three distribution models' weights, using ArcMap 10.5. To achieve the AUC value, IBM SPSS Statistics 19 was employed.

RESULTS

Climatic variables in the warm season (monthly average rainfall and monthly maximum average temperature) were highly correlated (-0.97). Therefore, the variable of monthly average rainfall in the warm season was eliminated due to the result of PCA. Meanwhile, in the cold season there was no correlation between data .

To validate the model performance, metric of area under Receiver Operating Characteristic (ROC) curve (AUC) was used (Table 1). According to the AUC values, RF and MaxEnt models performed great results (AUC > 0.9). Similarly, CTA models were also valid.

TABLE 1. AUC values of each model in both seasons.

	Random Forests	Classification Tree	Maximum Entropy
warm season	0.97	0.90	0.95
cold season	0.95	0.83	0.94

According to the results of all three models, the most important parameters of habitat suitability of Wild goat in the study area in both warm and cold seasons are altitude, distance to water resources, vegetation type, and slope. Among the existing plant species, Wild goat prefers *Artemisia sieberi*, *Stipa arabica*, *Zyghophylom eurpterum*, and *Sieditzia rozmarinus* in the warm season. Also, in the cold season in addition to the aforementioned plant species, *Dendrostellara lessertii* is another favorite plant for Wild goat. Meanwhile, with altitudes raising in the warm season, the suitability of the habitat grows. On the other hand, in cold season suitability does not increase with altitudes higher than 1400 m. Also, slopes above 20% and 35% respectively in warm and cold seasons and distances less than 5,000 meters to water resources are favorable to the species.

Habitat suitability map

The ensemble habitat suitability maps of Wild goat in Kavir National Park region based on weight-averaging of the three models in both seasons are illustrated in Figure 3. The main habitats of the species are the three mountain ranges labeled as A, B, C. The AUC values for both seasons were 0.93. The obtained areas as suitable habitats were slightly different in each model. In the ensemble model, the suitable areas in cold and warm season were 519.57 square kilometers (11% of the total area) and 581.05 square kilometers (13% of the total area), respectively.
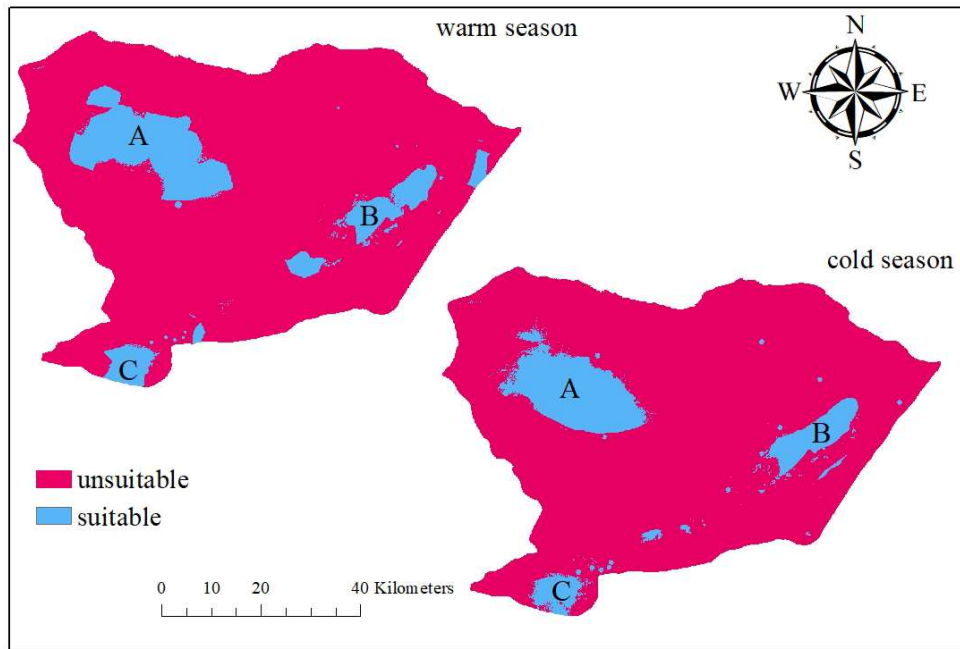
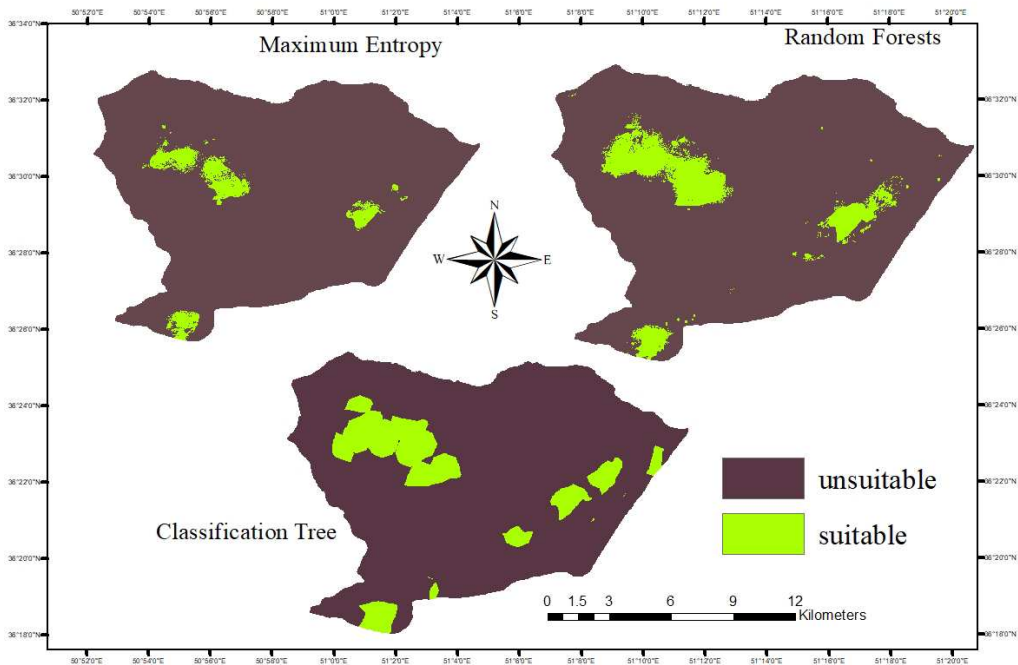**FIGURE 3.** Wild goat habitat suitability ensemble maps in the study area.

FIGURE 4. Binary maps of suitable area resulted by each model in warm season.
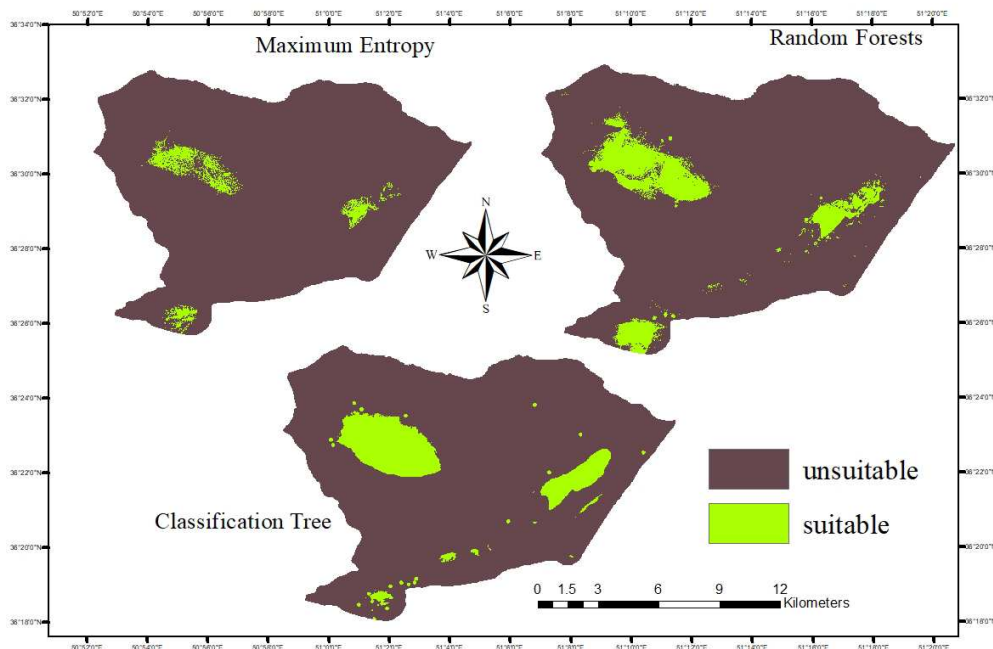
FIGURE 5. Binary maps of suitable area resulted by each model in cold season.

The results of each model are illustrated in the form of binary maps (Figs. 4, 5). In order to provide binary maps, baseline (SPM Users Guide Introducing CART, 2013; SPM Users Guide Introducing Random Forest, 2016) and 10 percentile training presence logistic (Young *et al.*, 2011) thresholds were used to distinguish suitable from unsuitable areas for the results of SPM and MaxEnt software.

DISSCUSSION
Suitable patches
Based on obtained data, it can be observed that apart from the labeled areas, there are some patches assigned to suitable category which are in great distances from the original habitats (Fig. 2). This suggests the possibility of Wild goat movement between their mainland, especially between B and C regions. It seems that a synergy of elevation, water resources, and vegetation type factors caused the larger patches (A, B, and C) to be suitable for Wild goats. Surprisingly, the smaller patches contain water resources at low altitudes. It means that in both seasons Wild goats are prone to pass plains in need of water. Occasionally, the species has been sighted in southern patches. Due to the existence of hunters and predators, movement of the species in plains can be unsafe.

Utilized models comparison
In this study, it was observed that the highest accuracy based on AUC in each season is obtained by the Random Forest model, and then, with a slight difference, the MaxEnt approach was presented with an excellent accuracy. In a paper by Williams (2009), a comparison was made between approaches of GLM, RF, ANN, and MaxEnt, and in conclusion, RF was known as the best model, followed by Maxnet with a small difference; which resembled the results of this study. Meanwhile, in a study by Ariel *et al.* (2017), the results of the RF algorithm, among GLM, MaxEnt, MARS, and BRT methods, were less accurate. In this study, both Classification Tree models performed less accurately than the others, which confirmed the results of Mi *et al.* (2017).

Habitat selection in the seasons

Based on the RF results, which was more accurate than the others, elevation is the most important predictor; In warm season, habitats with altitude more than 1050 meters above sea level is highly suitable on account of relative safety and lower temperature in high elevations. On the other hand, the optimum altitude in cold season is about 1,400 meters. It seems that low temperatures in this season may cause altitude above 1,400 m no more and even less preferable for the species. In a study in Yellowstone National Park, it was proven that in summer, higher altitudes are more preferable for Wild goats (DeVoe *et al.*, 2015). In the studies of Farashi *et al.* (2010) and Hosseini *et al.* (2016), altitudes have known to be important factors in determining the suitable habitats of this species.

Distance-to-water-resources is another factor that has a significant impact on the distribution of the species. The probability of species occurring at distances of more than 5000 meters to water resources is very low and close to zero. The significance of this variable in the cold season has declined in comparison to the warm season. This happens because Kavir's temperature in the warm season can easily reach 50 ° C, so in this condition, water would be a vital and crucial requirement. As Roberts (1977) indicated, lack of water resources in arid habitats is a limiting factor for many species. Also, in southeastern parts of the United States, water is known to be a limiting factor for ungulate species (Miller, 1984; Sa'nchez-Rojas, 2000).

Vegetation type is another influential factor. It seems that in the warm season, female goats demand more vegetation due to childbirth and lactation. On the other hand, foraging in a severe cold season is challenging. Therefore, vegetation variable plays an important role in Wild goat distribution. Also, in the study by Hosseini *et al.* (2016), vegetation is one of the influential factors in determining Wild goat favorite habitat.

In the present study, slopes over 20% and 35% are preferable for this species in warm and cold seasons, respectively. It seems that in the cold season, requiring relatively less water, makes Wild goat to choose habitats more freely and to draw more attention to steeper slopes. Several studies, including Moravati (2014), are in concordance with this study about the impact of slope. Wild goat is physically adapted to complex landscapes and steep slopes. This adaptation was originally developed as a strong strategy to reduce the risk of being hunted. It means that this species is able to quickly pass steep slopes (Adams, 1982; Fox & Streveler, 1986), while its natural predators such as leopard, despite its potential in crossing hills and steeps, cannot do it as skillfully as Wild goat can.

Conservation implementation

Kavir National Park climate is generally hot and arid. In this study, it was proved that water is one of the most important factors to determine habitat suitability. A number of water resources dry every year during summer, many of which are restored by game guards of the Department of Environment. But researchers have suggested that the development of water resources for wildlife, without study, might not have the expected benefits and even might cause negative impacts, because it can increase the amount of hunting, competition, and contagious diseases (Rosenstock *et al.* 1999). This is a serious issue, so studying water resources status is a prerequisite to any activity of restoring the habitat.

Due to the fact that there are some suitable patches between the mountains, it is suggested that further studies be conducted about corridors. On the other hand, hunters can ride their motorcycles in the plains, easily. Hence there is a need to secure the patches.

Another necessary implication is to teach tourism not to use the plant species as firewood. The density of vegetation in this region is relatively low and destroying this vegetation cover can directly cause adverse effects on Wild goats.

LITERATURE CITED

Adams, L.G., 1982. Ecological relationships of mountain goats and Rocky Mountain bighorn sheep. Proceedings of the 3rd Biennial Symposium. Northern Wild Sheep and Goat Council Colorado USA: Fort Collin 9–22.

Altman, A., Toloşi, L., Sander, O., Lengauer, T., 2010. Permutation importance: a corrected feature importance measure. Bioinformatics 26(10), 1340-1347.

Araújo, B., New, M., 2007. Ensemble forecasting of species distributions. Trends in ecology and evolution 22(1), 42-47.

Ariel, E.L., Valdez, R., Michel, S., 2017. Summer and winter habitat suitability of Marco Polo argali in southeastern Tajikistan: A modeling approach. Heliyon 3(11), e00445.

Atarodi, M., 2010. Asiatic Cheetah Habitats, Volume 1, National Park and protected area, Kavir. Tehran: Barg-e Zeitoon (Olive Leaf) 1, 23-26 (In Persian).

Breiman, L., 2001. Random Forest. Berkley: Statistics Department University of California, Berkeley, CA, 4720.

Breiman, L., Cutler, A., 2007. Random forests-classification description. Department of Statistics, Berkeley, 2.

Breiman, L., Friedman, J., Olshen, R., Stone, C., 1984. Classification and regression trees. Belmont: Wadsworth, International Group 432, 151-166

Brown J.L., 2014. SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic, and species distribution model analyses. Methods in Ecology and Evolution 5(7), 694-700.

Childs, C., 2004. Interpolating surfaces in ArcGIS spatial analyst. ArcUser, July-September, 3235, 569.

DeVoe, J.D., Garrott, R.A., Rotella, J.J., Challender, S.R., White, P.J., O'Reilly, M., Butler, C.J., 2015. Summer range occupancy modeling of non-native mountain goats in the greater Yellowstone area. Ecosphere 6(11), 1-20

Džeroski, S., 2001. Applications of symbolic machine learning to ecological modelling, Ecological Modelling 146, 263-276.

Elith, J., Phillips, S. J., Hastie, T., Dudı´k, M., Chee, Y.E., Yates, C. J., 2010. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17(1), 43- 57.

Farashi, A., Kaboli, M., Momeni, I., 2010. Habitat Suitability Modeling for Wild Goat Capra aegagrus in Kolah Ghazi National Park, Esfahan Province. Iran natural resources 62-73 (in Persian with English abstract).

Fook, K.D., Amaral, S., Monteiro, A. M., Câmara, G., Ximenes, A.D., Arasato, L.S., 2009. Making species distribution models available on the web for reuse in biodiversity experiments: Euterpe edulis species case study. Sociedade and Natureza 21(1), 39-49.

Fox, J.L., Streveler, G.P., 1986. Wolf predation on mountain goats in southeastern Alaska. Mammalogy 67, 192–195.

Gross, J.E., Kneeland, M.C., Reed, D.F., Reich, R.M., 2002. GIS-based habitat models for mountain goats. Journal of Mammalogy 83(1), 218-228.

Hirzel, A., Guisan, A., 2002. Which is the optimal sampling strategy for habitat suitability modelling. Ecological Modelling 157, 331-341.

Hosseini, S., Riazi, B., Shams Esfandabad, B., Naderi, M., 2016. Assessing habitat suitability for Wild goat in Golestan. Animal Environment 9(2), 9-16 (in Persian with English abstract).

Khosravi, R. Hemami, M.R., Malekian M., 2014. Comparison of Geostatistical Methods to Determine the Best Bioclimatic Data Interpolation Method for Modelling Species Distribution in Central Iran. Iranian Journal of Applied Ecology 3(8), 55-68.

Kocev, D., Naumoski, A., Mitreski, K., Krstić, S., Džeroski, S., 2010. Learning habitat models for the diatom community in Lake Prespa. Ecological Modelling 221(2), 330-337.

Kramer-Schadt, S.N., 2013. The importance of correcting for sampling bias in MaxEnt species distribution models. Diversity and Distributions 19(11), 1366-137.

Leitão, P., Moreira, F., Osborne, P., 2011. Effects of geographical data sampling bias on habitat models of species distributions: A case study with steppe birds in southern Portugal. International Journal of Geographical Information Science 25, 439–453.

Mi, C., Huettmann, F., Guo, Y., Han, X., Wen, L., 2017. Why choose Random Forest to predict rare species distribution with few samples in large undersampled areas? Three Asian crane species models provide supporting evidence. PeerJ 5, e2849.

Miller, G.D., Cochran, M.H., Smith, E.L. 1984. Nighttime activity of desert bighorn sheep. Desert Bighorn Council Transactions 28, 23-25.

Moravati, M. M., 2014. Desirable Areas and Effective Environmental Factors of Wild Goat Habitat (*Capra aegagrus*). International Journal of Environmental Research 8(4), 1031-1040.

Naderi, G., Saadati, H., Khalatbari, M., Imani, A., Riazi, B., 2014. Wild Bezoar Goat (*Capra aegagrus erxleben*) in Agh-Dagh Protected Area, Ardabil, Iran: Habitat Suitability Modelling by Ecological Niche Factor Analysis. Acta Zoologica Bulgarica 66(3), 385-388.

Naderi, S., Rezaei, H.R., Pompanon, F., Blum, M.G., Negrini, R., Naghash, H.R., Balkız, Ö., Mashkour, M., Gaggiotti, O.E., Ajmone-Marsan, P., Kence, A., 2008. The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. Proceedings of the National Academy of Sciences 105(46), 17659-17664.

Perry, M., Hollis, D., 2005. The generation of monthly gridded datasets for a range of climatic variables over the UK. International Journal of Climatology 25, 1041–1054.

Peterson, A., Vargas, N., 1993. Ecological diversity in Scrub Jays, *Aphelocoma coerulescens*. The Biological Diversity of Mexico: Origins and Distribution 309–317.

Phillips, S. J., Anderson, R., Schapire R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190, 231–259.

Roberts, R., 1977. Big game guzzlers. Rangeman's Journal 4(3), 80-82.

Rosenstock, S.S., Ballard, W.B., Devos, J.J., 1999. Benefits and impacts of wildlife water developments. Range Management 302-311.

Sa´nchez-Rojas, G.A., 2000. Mule deer (*Odocoileus hemionus*) density in a landscape element of the Chihuahuan Desert, Mexico. Arid Environments 44, 357–368.

Salford, SPM Users Guide Introducing CART. 2013. Retrieved from Salford-System.

Salford, SPM Users Guide Introducing Random Forest, 2016. Retrieved from Salford-System.

Sarhangzadeh, A., Yavari, M., Hemami, H., 2013. Habitat suitability modeling for wild goat (*Capra aegagrus*) in a mountainous arid area, central Iran. Caspian Journal of Environmental Sciences, 41-51.

Shams Esfandabad, B., Karami, M., Hemami, M.R., Riazi, B., Sadough, M.B., 2010. Habitat associations of wild goat in central Iran: implications for conservation. European Journal of Wildlife 56(6), 883-894.

Sluiter, R., 2009. Interpolation methods for climate data: literature review. KNMI intern rapport, Royal Netherlands Meteorological Institute, De Bilt.

Stockwell, D., Townsend Peterson, a., 2002. Effects of sample size on accuracy of species distribution models. Ecological Modelling 148(1), 1-13.

Taylor, S., Wall W., Kulis, Y., 2001. Habitat selection by mountain goats in south coastal British Columbia. In Biennial Symposium of the Northern Wild Sheep and Goat Council 15, 141-157.

Weinberg, P.J., 2008. *Capra aegagrus* (Bezoar, Wild Goat). Retrieved from The IUCN Red List of Threatened Species: http://www.iucnredlist.org/details/3786/0.

Williams, S.H., 2001. Multiple determinants of Australian tropical frog biodiversity. Biological Conservation 98(1), 1-10.

Young, N., Carter, L., Evangelista, P., 2011. A MaxEnt Model v3.3.3e Tutorial (ArcGIS v10). Natural Resource Ecology Laboratory at Colorado State University and the National Institute of Invasive Species Science.